DOCUMENT RESUME

ED 167 621 TM 008 435

AUTHOR Hollingsworth, Holly
TITLE The Coefficients of a Maximum Contrast as Interpretable Statistics.
PUB DATE Mar 78
NOTE 8p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Analysis of Variance; Memory; Primary Education; *Statistical Analysis; *Tests of Significance

ABSTRACT

A fundamental fact of the analysis of variance statistical procedure is that if the omnibus F test of an effect is significant, then there exists at least one contrast of that effect that will be significantly different from zero according to the S-method of Scheffe. The caveat to this rule is that the significant contrast(s) may not be of any interest to or interpretable by the investigator. Frustration does arise when the investigator determines the omnibus F test to be significant and then fails to find any meaningful group comparisons to be significant. The purpose of this paper is to discuss the coefficients of a contrast which will yield the largest test statistic according to the S-method. Then according to the fundamental theorem, this contrast will be significantly different from zero and, additionally, should offer the investigator some assistance in developing a parsimonious interpretation of the data. The S-method is a post-hoc method of comparing experimental groups and is often referred to as a data-snooping or sifting device. The techniques discussed here are also of the data-snooping species and may be employed to generate new hypotheses or to amend the existing theory of a practical problem. (Author/CTM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED167621

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE COEFFICIENTS OF A MAXIMUM CONTRAST AS INTERPRETABLE STATISTICS

By

Holly Hollingsworth

Saint Louis University

MARCH 1978

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Holly Hollingsworth

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Paper presented at the Annual Meeting of AERA, Toronto, March, 1978.

Printed in the United States of America

TM008 435

The Coefficients of a Maximum Contrast as Interpretable Statistics

Holly Hollingsworth*

Saint Louis University

Introduction

A fundamental fact of the analysis of variance statistical procedure is that if the omnibus F test of an effect is significant, then there exists at least one contrast of that effect that will be significantly different from zero according to the S-method of Scheffé (1959, p. 70). The caveat to this rule is that the significant contrast(s) may not be of any interest to or interpretable by the investigator. Frustration does arise when the investigator determines the omnibus F test to be significant and then fails to find any meaningful group comparisons (usually pairwise comparisons) to be significant. The purpose of this paper is to discuss the coefficients of a contrast which will yield the largest test statistic according to the S-method. Then according to the fundamental theorem, this contrast will be significantly different from zero and, additionally, should offer the investigator some assistance in developing a parsimonious interpretation of the data.

It is well known that the S-method is a post-hoc method of comparing experiemtal groups and is often referred to as a data-snooping or sifting device (Keppel, 1973, p. 93). The techniques to be discussed here are also of the data-snooping species and may be employed to generate new hypotheses or to amend the existing theory of a practical problem.

*paper presented at the Annual Meeting of the American Educational Research Association, March, 1978.

Method

To ease the discussion, let us consider a one-way analysis of variance situation with a treatment groups and s observations in each group. Let $\bar{A}_i$ and $\bar{T}$ be the mean observation in group i and the grand mean, respectively. A contrast among the a population means $\mu_1, \ldots, \mu_a$ is a linear combination

$\psi = \sum_{i=1}^{a} c_i \mu_i$ where $\sum_{i=1}^{a} c_i = 0$ and at least two of the c_i's are non-zero.

The unbiased estimate of ψ is $\hat{\psi} = \sum_{i=1}^{a} c_i \bar{A}_i$.

The statistic SS(A comp) = $s\left[\sum_{i=1}^{a} c_i \bar{A}_i\right]^2 / \sum_{i=1}^{a} c_i^2$, defined by Keppel (1973, p.98), can be thought as the sum of squares due to the comparison of an effect A defined by a contrast. SS(A comp) has a chi-square distribution with one degree of freedom and is independent of the statistic, mean square within groups, MS(W).

Since SS(A comp) is bounded from above by the sum of squares between groups, SS(A), (Keppel, 1973, p.109), we will define a maximum contrast, $\hat{\psi}_{max}$, to be a contrast for which SS(A comp) = SS(A). The existence of $\hat{\psi}_{max}$ is discussed by Scheffé (1959, p.71) and a set of coefficients of $\hat{\psi}_{max}$ is given by Winer (1971, p.176). A set of coefficients slightly different from Winer's will now be introduced.

One set of coefficients, $\{c_i\}$, of a maximum contrast is $\{\sqrt{s}\ (\bar{A}_i - \bar{T})/\sqrt{SS(A)}\}$. It is not necessary to show the mathematical derivation of these coefficients, for it is not a difficult task to verify that this set satisfy the requirements of being a contrast, as defined, and of maximizing the sum of squares due to the contrast. As a point of interest, one derivation of the coefficients would be similar to that of Scheffé (1969, p.118) where the notion of Lagrange multipliers is utilized.

The major difference between Winer's set and the above set of coefficients allows for a further interpretation of the data. For the set of coefficients presented in this paper, the square of a coefficient, c_i^2, is the proportion of the between groups' variation attributable to group i. This follows from the equality $c_i^2 = (\bar{A}_i - \bar{T})^2 / \sum_{i=1}^{a} (\bar{A}_i - \bar{T})^2$, i.e. c_i^2 is the ratio of a part to the whole. Note for the set, $\{c_i\}$, $\sum_{i=1}^{a} c_i^2 = 1$.

Example

In a study reported by Derevensky (1976), sixty kindergarten children were asked to perform a memory task of placing colored sticks in a specific order under six different instructional conditions. The subjects were instructed to either watch and/or listen to an experimenter manipulate the sticks and then replicate the manipulation starting with an identical or a different spatial arrangement of the sticks. The mean for each of the six conditions are given in Table 1. It is assumed that the design was balanced, i.e. s = 10.

Table 1

MEANS FOR EACH CONDITION AND THE COEFFICIENTS OF THE MAXIMUM CONTRAST

Condition-Spatial Arrangement	$\bar{X}$	c_i	c_i^2
1. Visual-only-identical	2.10	-.3706	.1374
2. Verbal-only-identical	2.90	.2224	.0495
3. Visual-verbal-identical	3.00	.2965	.0879
4. Visual-only-different	1.60	-.7412	.5495
5. Verbal-only-different	3.00	.2965	.0879
6. Visual-verbal-different	3.00	.2965	.0879
TOTAL	2.60		

A one-way analysis of variance of the data revealed a significant instructional condition effect. However, simple pairwise comparisons of the groups were all judged to be statistically non-significant at the .05 level using the S-method.

Table 2

ANALYSIS OF VARIANCE

Source	df	SS	MS	F	p
Conditions	5	18.20	3.64	3.77	0.05
Error	54	52.20	.97		
TOTAL	59	70.40			

The coefficients of $\hat{\psi}_{max}$ and their squares are reported in Table 1. An observation of these coefficients suggest that performance under the visual only instructional conditions (conditions 1 and 4) are inferior to the other instructional conditions. The same interpretation would be made on the basis of a subjective inspection of the means (Table 1). However, interpretations based on the coefficients of the maximum contrast allow for a natural progression to post-hoc data analytic procedures. For example, it would seem natural to test the post-hoc hypothesis, $.5(\mu_1 + \mu_4) - .25(\mu_2 + \mu_3 + \mu_5 + \mu_6) = 0$, which would be considered as contrasting or comparing the visual only conditions with the other four instructional conditions, and this hypothesis would be rejected at the .05 level of significance. One might also want to interpret the significantly different from zero ($p<.05$) contrast, which approximates the maximum contrast, $(.3\mu_1 + .7\mu_4) - .25(\mu_2 + \mu_3 + \mu_5 + \mu_6)$. This contrast suggests that the visual only conditions produce a different effect than the other four instructional conditions and, furthermore, the spatial arrangement does differentially effect the performance of the subjects within the visual only condition.

The observations under the instructional condition number 4 is accounting for nearly 55% ($c_4^2 = .5495$) of the between group variation as defined by SS(A). So that if the observations under group 4 were removed from the analysis,

the between groups sum of squares would be more than halved and the omnibus F test may no longer be significant at the desired level of significance.

Discussion

One advantage of post-hoc comparisons is that the experimenter may uncover an interesting, but unplanned, comparison which could be incorporated into the existing theoretical framework of the research area and used to generate further hypotheses. In this respect, it is obligatory that the experimenter analyze the maximum contrast. New research, then can be designed to test hypotheses suggested by the maximum contrast and existing theory can be modified to reflect an observed state of nature.

When the experiment has unequal sample sizes, then the defining formula for the sum of squares between groups will differ depending on the method of analysis (Keppel, 1973, pp. 346-47). If one is using the unweighted means analysis, then the coefficients of $\hat{\psi}_{max}$ would be $\{\sqrt{\tilde{s}}(\bar{A}_i - \bar{T})/\sqrt{SS(A)}\}$ where $\tilde{s}$ is the harmonic mean of the sample sizes of the groups and $\bar{T} = \Sigma\bar{A}_i/a$.

A generalization of this discussion of the maximum contrast can be made to factorial designs. The reader is referred to Chapters 11 and 12 of Keppel (1973) for a treatment of multiple comparisons in a factorial experiment with two factors. It is natural to ask if a maximum contrast exists that could be used to estimate the degree to which an experimental cell contributes to the interaction of two or more factors. The answer is affirmative and that contrast will be discussed at a later date.

Finally, a comment about generalizability of the method is in order. The coefficients of a maximum contrast are computed from the observations

made on a sample. At best, these coefficients should be considered the same as regression coefficients with respect to their ability to estimate population characteristics. To avoid the errors of over interpretating the results of an experiment, it is suggested that new data be collected to validate, using a-priori comparison procedures, the interpretations based on a maximum contrast.

REFERENCES

Derevensky, J., Developmental changes in the effect of verval, nonverbal, and spatial-positional cues for memory. The Journal of Experimental Education, 1976, 45, 52-60.

Keppel, G., Design and Analysis: A Researchers Handbook, Englewood Cliffs: Prentice Hall, Inc., 1973.

Scheffé, H., The Analysis of Variance, New York: John Wiley & Sons, Inc., 1959.

Winer, B.J., Statistical Principles in Experimental Design, New York: McGraw-Hill, 1971.